



PCT

## RAW SEQUENCE LISTING

DATE: 08/31/2004

PATENT APPLICATION: US/10/505,405

TIME: 16:45:27

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\08312004\J505405.raw

3 <110> APPLICANT: COLE, Stewart  
 4       BROSCH, Roland  
 5       GORDON, Stephen  
 6       EIGLMEIER, Karin  
 7       GARNIER, Thierry  
 9 <120> TITLE OF INVENTION: DELETED SEQUENCE IN M. TUBERCULOSIS, METHOD FOR  
 10       DETECTING MYCOBACTERIA USING THESE SEQUENCES AND  
 11       VACCINES  
 13 <130> FILE REFERENCE: 346 379- US  
 C--> 15 <140> CURRENT APPLICATION NUMBER: US/10/505,405  
 C--> 15 <141> CURRENT FILING DATE: 2004-08-24  
 15 <150> PRIOR APPLICATION NUMBER: PCT/IB 03/00 986  
 16 <151> PRIOR FILING DATE: 2003-02-25  
 18 <150> PRIOR APPLICATION NUMBER: EP 02/290 458  
 19 <151> PRIOR FILING DATE: 2002-02-25  
 21 <160> NUMBER OF SEQ ID NOS: 22  
 23 <170> SOFTWARE: PatentIn Ver. 2.1

**Does Not Comply  
Corrected Diskette Needed**

## ERRORED SEQUENCES

1879 <210> SEQ ID NO: 21  
 1880 <211> LENGTH: 1191  
 1881 <212> TYPE: DNA  
 1882 <213> ORGANISM: Mycobacterium tuberculosis  
 1884 <220> FEATURE:  
 1885 <221> NAME/KEY: CDS  
 1886 <222> LOCATION: (1)..(1191)  
 1887 <223> OTHER INFORMATION: Fusion gene between mmpS6 and mmpL6 genes  
 1889 <220> FEATURE:  
 1890 <221> NAME/KEY: misc\_feature  
 1891 <222> LOCATION: (1) (1191)  
 1892 <223> OTHER INFORMATION: CDS corresponds to fusion protein of rearranged forms  
 1893       of mmpS6 and mmpL6  
 E--> 1895 <400> SEQUENCE: ①21  
 1896 gtg cag ggg att tca gtg act ggc ctg gtc aaa cgc ggc tgg atg gtg   48  
 1897 Val Gln Gly Ile Ser Val Thr Gly Leu Val Lys Arg Gly Trp Met Val  
 1898   1                   5                   10                   15  
 1900 aga tcc gtc ttt gac acg atc gac ggt atc gac caa ctc ggc gag cag   96  
 1901 Arg Ser Val Phe Asp Thr Ile Asp Gly Ile Asp Gln Leu Gly Glu Gln  
 1902                   20                   25                   30  
 1904 ctg gcc agc gtg acc gta acc ttg gac aag ttg gct gcg atc cag cct   144  
 1905 Leu Ala Ser Val Thr Val Thr Leu Asp Lys Leu Ala Ala Ile Gln Pro

## RAW SEQUENCE LISTING

DATE: 08/31/2004

PATENT APPLICATION: US/10/505,405

TIME: 16:45:27

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\08312004\J505405.raw

1906	35	40	45	
1908	caa ttg gtg gcg ctg cta cca gac gag atc gcc agc cag cag atc aat	192		
1909	Gln Leu Val Ala Leu Leu Pro Asp Glu Ile Ala Ser Gln Gln Ile Asn			
1910	50 55 60			
1912	cgg gaa ctg gcg ctg gct aac tac gcc acc atg tcc ggg atc tat gcc	240		
1914	Arg Glu Leu Ala Leu Ala Asn Tyr Ala Thr Met Ser Gly Ile Tyr Ala			
1915	65 70 75 80			
1917	cag acg gcg gcc ttg atc gaa aac gct gcc gcc atg gga caa gcc ttt	288		
1918	Gln Thr Ala Ala Leu Ile Glu Asn Ala Ala Ala Met Gly Gln Ala Phe			
1919	85 90 95			
1921	gac gcc gcc aag aac gac gac tcc ttc tat ctg ccg ccg gag gct ttt	336		
1922	Asp Ala Ala Lys Asn Asp Asp Ser Phe Tyr Leu Pro Pro Glu Ala Phe			
1923	100 105 110			
1925	gac aac cca gat ttc cag cgc ggc ctg aaa ttg ttc ctg tcg gca gac	384		
1926	Asp Asn Pro Asp Phe Gln Arg Gly Leu Lys Leu Phe Leu Ser Ala Asp			
1927	115 120 125			
1929	ggt aag gcg gct cgg atg atc atc tcc cat gaa ggc gat ccc gcc acc	432		
1930	Gly Lys Ala Ala Arg Met Ile Ile Ser His Glu Gly Asp Pro Ala Thr			
1931	130 135 140			
1933	ccc gaa ggc att tcg cat atc gac gcg atc aag cag gcg gcc cac gag	480		
1934	Pro Glu Gly Ile Ser His Ile Asp Ala Ile Lys Gln Ala Ala His Glu			
1935	145 150 155 160			
1937	gcc gtg aag ggc act ccc atg gcg ggt gct ggg atc tat ctg gcc ggc	528		
1938	Ala Val Lys Gly Thr Pro Met Ala Gly Ala Gly Ile Tyr Leu Ala Gly			
1939	165 170 175			
1941	acg gcc gcc acc ttc aag gac att caa gac ggc gcc acc tac gac ctc	576		
1942	Thr Ala Ala Thr Phe Lys Asp Ile Gln Asp Gly Ala Thr Tyr Asp Leu			
1943	180 185 190			
1945	ctg atc gcc gga ata gcc gcg ctg agc ttg att ttg ctc atc atg atg	624		
1946	Leu Ile Ala Gly Ile Ala Ala Leu Ser Leu Ile Leu Leu Ile Met Met			
1947	195 200 205			
1949	atc att acc cga agc ctg gtt gcg gcg ctg gtg atc gtg ggc acg gtg	672		
1950	Ile Ile Thr Arg Ser Leu Val Ala Ala Leu Val Ile Val Gly Thr Val			
1951	210 215 220			
1953	gcg ctg tcg ttg ggc gct tct ttt ggc ctg tcc gtg ctg gtg tgg cag	720		
1954	Ala Leu Ser Leu Gly Ala Ser Phe Gly Leu Ser Val Leu Val Trp Gln			
1955	225 230 235 240			
1957	cat ctt ctc ggt atc cag ttg tac tgg atc gtg ctc gcg ctg gcc gtc	768		
1958	His Leu Leu Gly Ile Gln Leu Tyr Trp Ile Val Leu Ala Leu Ala Val			
1959	245 250 255			
1961	atc ctg ctc ctg gcc gtg gga tcg gac tat aac ttg ctg ctg att tcc	816		
1962	Ile Leu Leu Leu Ala Val Gly Ser Asp Tyr Asn Leu Leu Leu Ile Ser			
1963	260 265 270			
1965	cga ttc aag gag gag atc ggt gca ggt ttg aac acc ggc atc atc cgt	864		
1966	Arg Phe Lys Glu Glu Ile Gly Ala Gly Leu Asn Thr Gly Ile Ile Arg			
1967	275 280 285			
1969	gcg atg gcc ggc acc ggc ggg gtg gtg acc gct gcc ggc ctg gtg ttc	912		
1970	Ala Met Ala Gly Thr Gly Gly Val Val Thr Ala Ala Gly Leu Val Phe			
1971	290 295 300			

## RAW SEQUENCE LISTING

DATE: 08/31/2004

PATENT APPLICATION: US/10/505,405

TIME: 16:45:27

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\08312004\J505405.raw

```

1973 gcc gcc act atg tct tcg ttc gtg ttc agt gat ttg cgg gtc ctc ggt      960
1974 Ala Ala Thr Met Ser Ser Phe Val Phe Ser Asp Leu Arg Val Leu Gly
1975 305                               310                               315                               320
1977 cag atc ggg acc acc att ggt ctt ggg ctg ctg ttc gac acg ctg gtg      1008
1978 Gln Ile Gly Thr Thr Ile Gly Leu Gly Leu Leu Phe Asp Thr Leu Val
1979                               325                               330                               335
1981 gtg cgc gcg ttc atg acc ccg tcc atc gcg gtg ctg ctc ggg cgc tgg      1056
1982 Val Arg Ala Phe Met Thr Pro Ser Ile Ala Val Leu Leu Gly Arg Trp
1983                               340                               345                               350
1985 ttc tgg tgg ccg caa cga gtg cgc ccg cgc cct gcc agc agg atg ctt      1104
1986 Phe Trp Trp Pro Gln Arg Val Arg Pro Arg Pro Ala Ser Arg Met Leu
1987                               355                               360                               365
1990 cgg ccg tac ggc ccg cgg ccc gtg gtt cgt gaa ttg ctg ctg cgc gag      1152
1991 Arg Pro Tyr Gly Pro Arg Pro Val Val Arg Glu Leu Leu Leu Arg Glu
1992                               370                               375                               380
1994 ggc aac gat gac ccg aga act cag gtg gct acc cac cgt                        1191
1995 Gly Asn Asp Asp Pro Arg Thr Gln Val Ala Thr His Arg
1996 385                               390                               395
1999 <210> SEQ ID NO: 22
2000 <211> LENGTH: 397
2001 <212> TYPE: PRT
2002 <213> ORGANISM: Mycobacterium tuberculosis
2004 <220> FEATURE:
2005 <223> OTHER INFORMATION: Fusion protein of rearranged forms of mmpS6 and mmpL6
E--> 2007 <400> SEQUENCE: ② 22
2008 Val Gln Gly Ile Ser Val Thr Gly Leu Val Lys Arg Gly Trp Met Val
2009 1                               5                               10                               15
2011 Arg Ser Val Phe Asp Thr Ile Asp Gly Ile Asp Gln Leu Gly Glu Gln
2012                               20                               25                               30
2014 Leu Ala Ser Val Thr Val Thr Leu Asp Lys Leu Ala Ala Ile Gln Pro
2015                               35                               40                               45
2017 Gln Leu Val Ala Leu Leu Pro Asp Glu Ile Ala Ser Gln Gln Ile Asn
2018                               50                               55                               60
2020 Arg Glu Leu Ala Leu Ala Asn Tyr Ala Thr Met Ser Gly Ile Tyr Ala
2021 65                               70                               75                               80
2023 Gln Thr Ala Ala Leu Ile Glu Asn Ala Ala Ala Met Gly Gln Ala Phe
2024                               85                               90                               95
2027 Asp Ala Ala Lys Asn Asp Asp Ser Phe Tyr Leu Pro Pro Glu Ala Phe
2028                               100                              105                              110
2030 Asp Asn Pro Asp Phe Gln Arg Gly Leu Lys Leu Phe Leu Ser Ala Asp
2031                               115                              120                              125
2033 Gly Lys Ala Ala Arg Met Ile Ile Ser His Glu Gly Asp Pro Ala Thr
2034                               130                              135                              140
2036 Pro Glu Gly Ile Ser His Ile Asp Ala Ile Lys Gln Ala Ala His Glu
2037 145                              150                              155                              160
2039 Ala Val Lys Gly Thr Pro Met Ala Gly Ala Gly Ile Tyr Leu Ala Gly
2040                               165                              170                              175
2042 Thr Ala Ala Thr Phe Lys Asp Ile Gln Asp Gly Ala Thr Tyr Asp Leu
2043                               180                              185                              190

```

## RAW SEQUENCE LISTING

DATE: 08/31/2004

PATENT APPLICATION: US/10/505,405

TIME: 16:45:27

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\08312004\J505405.raw

```

2045 Leu Ile Ala Gly Ile Ala Ala Leu Ser Leu Ile Leu Leu Ile Met Met
2046           195                      200                      205
2048 Ile Ile Thr Arg Ser Leu Val Ala Ala Leu Val Ile Val Gly Thr Val
2049       210                      215                      220
2051 Ala Leu Ser Leu Gly Ala Ser Phe Gly Leu Ser Val Leu Val Trp Gln
2052 225                      230                      235                      240
2054 His Leu Leu Gly Ile Gln Leu Tyr Trp Ile Val Leu Ala Leu Ala Val
2055           245                      250                      255
2057 Ile Leu Leu Leu Ala Val Gly Ser Asp Tyr Asn Leu Leu Leu Ile Ser
2058           260                      265                      270
2060 Arg Phe Lys Glu Glu Ile Gly Ala Gly Leu Asn Thr Gly Ile Ile Arg
2061           275                      280                      285
2063 Ala Met Ala Gly Thr Gly Gly Val Val Thr Ala Ala Gly Leu Val Phe
2064       290                      295                      300
2066 Ala Ala Thr Met Ser Ser Phe Val Phe Ser Asp Leu Arg Val Leu Gly
2067 305                      310                      315                      320
2069 Gln Ile Gly Thr Thr Ile Gly Leu Gly Leu Leu Phe Asp Thr Leu Val
2070           325                      330                      335
2072 Val Arg Ala Phe Met Thr Pro Ser Ile Ala Val Leu Leu Gly Arg Trp
2073           340                      345                      350
2075 Phe Trp Trp Pro Gln Arg Val Arg Pro Arg Pro Ala Ser Arg Met Leu
2076           355                      360                      365
2078 Arg Pro Tyr Gly Pro Arg Pro Val Val Arg Glu Leu Leu Leu Arg Glu
2079       370                      375                      380
2081 Gly Asn Asp Asp Pro Arg Thr Gln Val Ala Thr His Arg
2082 385                      390                      395

```

E--> 2089 (16) delete

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/505,405

DATE: 08/31/2004

TIME: 16:45:28

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\08312004\J505405.raw

L:15 M:270 C: Current Application Number differs, Replaced Current Application No

L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:1895 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:21 differs:1

L:2007 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:22 differs:2

L:2089 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:2